

Northumbria Research Link

Citation: Ellman, Jeremy, Lee, Nathan and Jin, Nanlin (2018) Cloud Computing Deployment: A Cost-Modelling Case-Study. Wireless Networks. ISSN 1022-0038 (In Press)

Published by: Springer

URL: <https://doi.org/10.1007/s11276-018-1881-2> <<https://doi.org/10.1007/s11276-018-1881-2>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/35210/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary



Cloud computing deployment: a cost-modelling case-study

Jeremy Ellman¹ · Nathan Lee¹ · Nanlin Jin¹

© The Author(s) 2018

Abstract

There are a wide range of cloud services commercially available. However there is limited research that investigates the strengths and weaknesses of their cost models in relation to different types of usage requirements. We propose a new costing model that systematically evaluates cloud services, and which combines compute, disk storage, and memory requirements. This paper demonstrates the proposed costing model on a data set that was derived from a real-world industrial data centre workload by calculating the precise cost of service provision from two leading cloud providers.

Keywords Cloud computing costs · Industrial IT workload

1 Introduction

1.1 Background

Currently many organizations are embracing dynamic, cloud-based operating models to position themselves for cost optimization and increased competitiveness. To address this, we propose a new costing model that measures the level of different cloud services required.

Cloud computing refers to the use of shared computing resources [2]. It may also be characterized as a pay-per-use model for enabling available, on-demand network access to a shared pool of configurable computing resources that can rapidly be provisioned with minimal management effort or service provider interaction [10]. A Computing Cloud can be hosted either privately, or publicly. [11] describes Public Cloud as hosted at the vendor's premises, giving the customer no visibility over the location of the cloud infrastructure. In both private and public Clouds, the three main services are Infrastructure as a Service, Platform as a Service and Software as a Service [5].

Infrastructure as a Service (IaaS) is the delivery of hardware (server, storage and network) and virtual operating systems as a service. IaaS provider does very little management other than keep the data centre operational and relies on the client to be able to manage the software services, as they would in their own data centre [2]. Platform as a Service (PaaS) is an infrastructure with applications supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the applications and configuration for the application-hosting environment [7]. Software as a Service (SaaS) is the capability to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices. The consumer does not manage or control the infrastructure or platform but may have access to software configuration [7]. These three terms are widely used across all the main cloud providers [6].

The costs of hybrid cloud models differ from provider to provider. Some charge for compute resources per minute and others by month. Others too, offer multiple services with combinations of charging schemes. Consequently the contribution of this paper is the creation of cost equations that allow meaningful comparison between cloud providers in the context of an Enterprise IT workload.

Research [3] states that the leaders in the private cloud industry are Amazon Web Services, Microsoft Azure, IBM, Google, and Salesforce. In 2015 Amazon Web Services share of the worldwide market was 31%, followed

✉ Jeremy Ellman
jeremy.ellman@northumbria.ac.uk
Nanlin Jin
nanlin.jin@northumbria.ac.uk

¹ Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne, UK

by Microsoft (9%), Google (4%.) The Cloud providers chosen for this work are Amazon Web Services (AWS), Microsoft's Azure, and Google Cloud Platform (GCP) as the largest providers of enterprise cloud services. All the suppliers offer an increasing range of more than fifty services each. However in this initial study only basic Compute and storage services are considered.

A major engineering manufacturer is currently undergoing a significant change in its European Enterprise IT Infrastructure. The current model consists of a data centre which houses 120 physical servers run either applications or a virtual private-cloud infrastructure. The Server and Storage team use a 'buy as required' model, which allows the data centre to host all the manufacturers' European servers and allows a small buffer for future growth.

There are two major issues with this model: Firstly keeping up-to-date with continuous improvement in current technologies, the Server and Storage Team are required to recycle physical equipment every three years. Secondly, it is challenging to monitor resource usage and to know when to increase capacity. Cloud Computing is attractive to business as it eliminates the requirement for forecasts and planning ahead of provisioning. It also permits companies to increase resource only when there is a rise in demand [13].

Organizations that embrace dynamic, cloud-based operating models position themselves better for cost optimization and increased competitiveness [9]. Consequently, the manufacturer plans to move their model to a public hybrid-cloud.

The paper proceeds as follows. Firstly the costs involved generally in the enterprise data centre are described. Next the specific details of size, usage, and capacity of one enterprise data centre are discussed. After an initial preamble on cloud service costs, cost modelling equations are developed that are designed to generalize over multiple cloud providers. A worked example of these equations is then applied to the actual data centre in the case study. Finally we conclude with a discussion of limitations and proposed future work.

2 Enterprise data centres and cloud providers

In this section the enterprise data centre used in this case study is firstly described, then appropriate cloud providers are considered.

2.1 Enterprise data centre costs

In a traditional data centre costs came from five main elements. These are the fixed costs of constructing the data

centre building and its power and cooling infrastructure, plus variable costs that increase with the level of output. These include the variable cost of populating the data centre with hardware, and finally the variable cost of operating and managing the servers in use [4].

In cloud computing, expenditure may be significantly reduced as fixed costs are removed and factored into the cloud vendors' pay per use pricing, where 'Pay per use', is an approach for pricing that allows customers to pay only for the individual services needed, and without requiring long-term contracts or licensing [1].

In a cloud environment, the variable costs would change each month depending on usage. Semi-variable or mixed costs have attributes of both fixed and variable costs. The introduction of cloud computing has also introduced a shift with cost models when compared to a traditional data centre. It is said that in the past, the vast majority of IT investments were either labour, or fixed costs. However cloud is removing this traditional fixed cost and replacing it with a variable cost component [12].

2.2 The enterprise IT data centre usage

In this section we look at the composition and usage of an Enterprise IT data centre. The centre is owned and operated by a large engineering manufacturing plant and runs a variety of business applications including enterprise resource planning, database and file servers, customer relationship management applications, etc.

Usage information was collected using IBM Tivoli Monitoring during one month where readings were taken every hour covering all the servers in the data centre. The information collected included, the servers' name, number of processors, percentage CPU usage, available and used RAM, and hard disk.

2.2.1 CPU

The data centre currently has 2879 CPUs across their full infrastructure of 504 servers, ($\mu = 5.7$) CPUs per server. During one month the mean CPU usage was 2.7%. Calculating this potential wastage aggressively could see a drop in the total required CPUs from 2879 to 78, and the averaged 5.7 CPU's per server decreased to 0.1539 CPU (rounded up to 1 CPU). The CPU usage standard deviation, $\sigma = 5.1\%$. However, peak usage was 82.1%. Consequently a local data centre needs to resource $\approx 10\times$ more capacity than needed 75% of the time.

2.2.2 Memory (RAM)

Currently the data centre uses 214,213 MB (210 GB) RAM across the 504 servers, averaging 425 MB per server. Of

this, mean usage is 43.1%. However, were memory to be rationed, applications could revert to using considerably slower virtual (disk based) memory. Such swapping would considerably affect performance, so no plan is offered to optimise ram usage.

2.2.3 Disk

The data centre currently has 476 TB of disk based storage across the 504 servers, giving a mean available disk size of 944,470 MB (944.47 GB) per server. However, disk size is highly right skewed (See Fig. 4), with a median (most common) disk volume size of 175 GB, whilst 75% disk volumes have < 500 GB. Nevertheless, a very small number of machines are file servers with more than 8 Tb storage.

Although the data centre uses mixed RAID/non-RAID systems, the proportions are not recorded. Furthermore, although all the cloud service providers do support RAID configurations, for simplicity only non-RAID systems are considered in this paper.

During one month overall disk storage had a mean utilization of 50.4%. Calculating this potential wastage aggressively could reduce the total required storage from 476,013,073 to 239,899,954 MB, and the mean 944,470 MBs per server decreased to 475,068 MBs per server.

The standard deviation for disk usage is 23.5%, with peak usage at 98.8%. A conservative model would see the cloud usage being based on the highest peak percentage but would still have reductions; the total storage usage would be reduced from 476,013,073 to 470,321,900 MB and see the 944,470 MB per server decreased to 933,136 MB.

2.2.4 Server classification

The servers in the data centre may be grouped by the number of CPU cores and RAM size. Several machines are file servers for the enterprise and have large disk capacity. However, disk volumes are readily interchanged and also easily adjusted in size. Consequently, servers will be classified here by number of CPU cores, and RAM only.

There are 41 different server types in the data centre, with varying storage capacities. These range from single core machines with 1 GB RAM, to 16 processor machines with 160 GB RAM, with the most common server type being dual core machines with 4 GB RAM. Table 1 shows the most popular machine configurations.

2.2.5 Activity

As can be seen from the summary chart, Fig. 1, overall server activity shows distinct cycles that likely relate to the

Table 1 Server classification by cores, and RAM with frequency

Server class	Cores	Ram (GB)	Count
S1	1	4	30
S2	2	4	79
S3	2	8	63
S4	2	16	37
S5	4	8	27
S6	4	12	23
S7	4	16	25
S8	4	32	24
S9	8	32	22
S10	12	160	6

working day. Individual servers may be heavily or lightly used, where they are active for 20% of the time or less. This is important, since different cloud cost models apply may be used for both groups (Figs. 2, 3).

2.3 Cost models implemented by cloud providers

2.3.1 Server tiers and size costs

Amazon Web Services' EC2 service provides a wide selection of server types. These consist of combinations of CPU, memory, storage, and networking capacity. Each server type includes one or more server sizes, allowing resources scaling [1]. Amazon Web Services offer 45 variations of Windows Instances ranging from 1 vCPU, 512 MB Memory and Elastic Block Storage (EBS) only to 349 vCPU, 1952 Memory and 3.75 TB SSD Storage. Amazon Web Services also offer 44 Linux Instances from 1 vCPU, 1 GB Memory, Memory and Elastic Block Storage (EBS) only to 349 vCPU, 1952 Memory and 3.75 TB SSD Storage (Figs. 4, 5, 6).

Azure [8] also offers a variety of servers on both Windows and multiple Linux variants. Comparing this with Amazon Web Services Linux Instances [1], AWS offer a wider range of Linux Instances, whilst Google offer a two-part server tiers for both Windows and Linux. The main tier sizes consist of the computing resource that is provided by Google, which is a similar to Amazon Web Services and Microsoft's Azure. However, Google adds an additional cost that depends on the operating system image loaded onto the server. (\$0.06 for Red Hat Linux Server with four or fewer CPUs, \$0.13 for Red Hat Linux Server with more than four CPUs).

All the Cloud providers use the concept of a Reserved Instance. Here, a VM is paid in advance for 1-year's (or longer) continual operation in exchange for a substantial

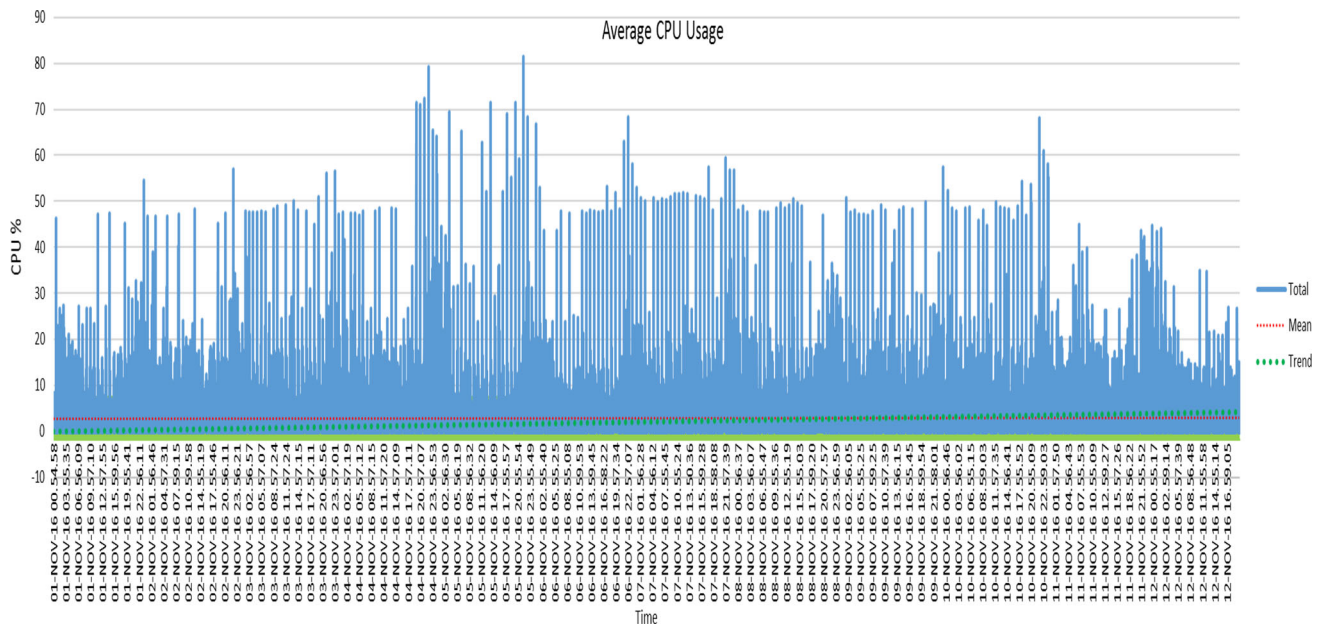


Fig. 1 Daily CPU usage versus time (1 month)

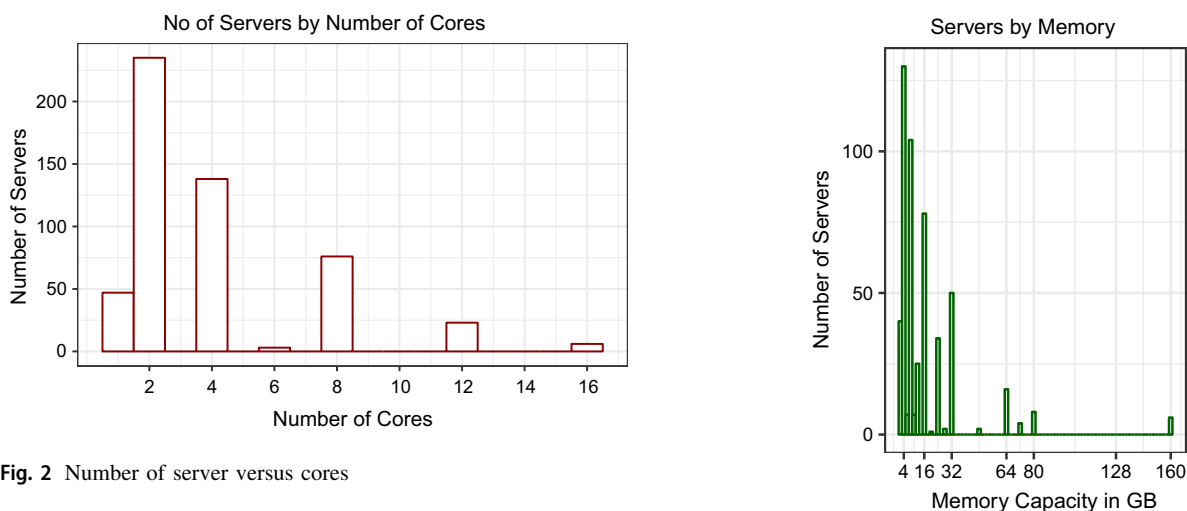


Fig. 2 Number of server versus cores

Fig. 3 # Server versus Ram (GB)

discount. This can significantly reduce virtual machine costs. For example, if a provider offers a 50% price reduction for a VM instance, and that machine is in operation for $> 50\%$ of the time, then it is more economic to pay in advance and operate that machine continually.

For Networking, Microsoft Azure offer four types of deployments: Public IP Address (Static or Dynamic), Cloud Service VIP, Reserved IP Address and Instance-Level Public IP Address (ILPIP). However, this is a different model to that of AWS, since AWS do not charge for IP addresses unless it has been reserved but is not attached to a running instance. This could be seen as a resource wastage fee. As with AWS, GCP do not charge for IP addresses unless an IP address has been allocated to an unused server.

3 A cloud solution

In this section we firstly choose several well known cloud service providers for comparison purposes. Next, a cost model is constructed that may be applied to multiple providers. This takes account of data centre usage patterns.

3.1 Picking cloud providers

Three cloud providers were needed for this initial study. The basis for the selection was general market presence and open cost disclosure. Amazon Web Services (aws.amazon.com) share of the worldwide market was

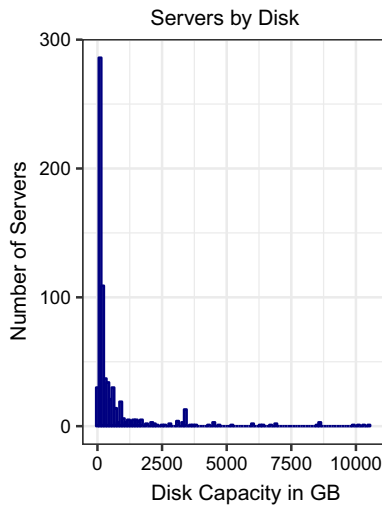


Fig. 4 # Server versus disk capacity (GB)

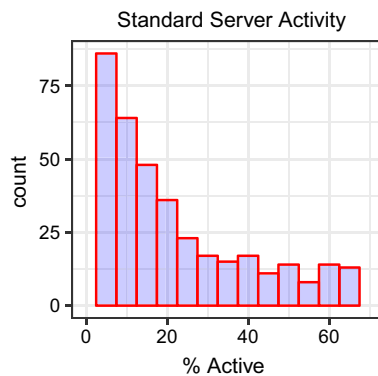


Fig. 5 # Standard servers versus % activity

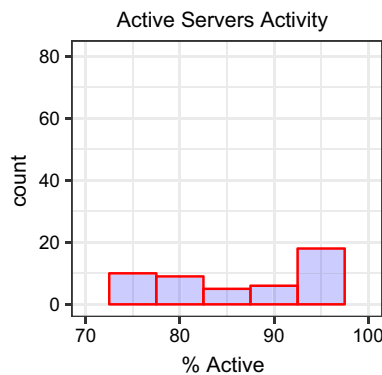


Fig. 6 # Active servers versus % activity

31%, followed by Microsoft (azure.microsoft.com) which has 9% and Google which has 4%. These three providers all offer on-line cost calculators too, hence these were selected. There are many other providers such as IBM, Oracle, Dell, Salesforce, and VMware who are either enterprise focused, or offer a more limited range of services, possibly where costs are not readily disclosed.

3.2 Cost models

Cloud Computing cost models become complex due to the variety of services each provider offers. To simplify this issue, we will focus on server and storage infrastructure; vCPU, vRAM, Storage (Disk), IP Allocation, Support Packages and Operating System. Amazon Web Services' EC2 service, Microsoft Azure and Google Cloud Platform provides a wide selection of server types designed to fit different use cases. Server types consist of combinations of CPU, memory, storage, and networking capacity and give the client the flexibility to choose the appropriate mix of resource [8, 1]. The method chosen will consist of analyzing this information from each of these providers to create a model that can be used to determine the most cost effective provider for any individual organization.

An alternative would be to use the cost calculators from each of providers to create an equation. These do not though take into account usage patterns over a range of devices. Additionally, each of the service providers calculate different levels of detail, which complicates a true comparison.

4 Our cost data calculations

If this section we propose a set of formulae that allow service costs to be compared between different cloud providers. The formulae take into account the actual server mix in a case study enterprise data centre, and their relative degree of activity.

In this section the subscripts A, M, G , are used to indicate AWS, Microsoft Azure, and Google Cloud Platform respectively, whilst T, C, D, I, S are defined as follows:

T	Total price
C	Compute cost
D	Cost of Persistent storage
I	IP Address price
S	Support cost

Then, the total cost per month is given by:

$$T = C + M + I + S \quad (1)$$

Now we go on to look at how each element may be calculated.

4.1 Compute pricing (C)

The principal components of compute costs are those of reserved, and on-demand instances. Let these be C_R , and

C_S respectively. Standard prices apply to both instances, although reserved instances attract a discount.

Given that a range of instances is required, their prices per hour may be stored in a vector. We define the pricing vector, \mathbf{p} as follows:

$$\mathbf{p} = [p_1, p_2, \dots, p_n] \quad (2)$$

where p_1, \dots, p_n are the hourly prices of the server classes equivalent to s_1, s_2, \dots, s_n from (1). We also define the frequency vector \mathbf{f} as:

$$\mathbf{f} = [f_1, f_2, \dots, f_n] \quad (3)$$

where f_1, \dots, f_n are the number of instances the server classes equivalent to s_1, s_2, \dots, s_n from (1) that are active.

Now, by definition, a reserved instance R , operates full time, which is 730 h per month. Let the frequency vector for the reserved instances then be \mathbf{f}_R . If the suppliers discount rate is D , then the cost per month C_R is:

$$C_R = 730(1 - D) \times \text{sum}(\mathbf{p} \cdot \mathbf{f}_R) \quad (4)$$

$$= 730(1 - D) \times \sum_{i=1}^n p_i f_i \quad (5)$$

Non-reserved instance costs are slightly more complex, since we need to find the product of the pricing vector \mathbf{p} , with the hours used for each server of a particular type. Here we require the frequency vector for standard (i.e. pay as you use) instances \mathbf{f}_S for each hour of the month. Let this be $\mathbf{f}_{h,d}$, where h, d correspond to the hour of the day, and the day of the month respectively. Then, the array $\mathbf{A}_{h,d}$ holds the frequency vectors $\mathbf{f}_{h,d}$ showing number of servers of each type in use during the hour h of day d .

$$C_S = \text{sum}(\mathbf{p} \cdot \mathbf{A}_{h,d}) \quad (6)$$

$$= \text{sum} \left(\mathbf{p} \cdot \sum_{h=1}^{23} \sum_{d=1}^{31} \mathbf{f}_{h,d} \right) \quad (7)$$

$$= \sum_{i=1}^n \sum_{h=1}^{23} \sum_{d=1}^{31} p_i f_{h,d} \quad (8)$$

4.2 Disk storage pricing (D)

Let D_A, D_M, D_G be the monthly persistence storage price per GB for the respective service providers (p). Since increased persistent storage is readily available, it is not necessary to provision the current maximum storage available in the data centre. Rather, we provide current storage used U_D , plus a margin for ready expansion M .

$$D = (U_D + M(U_D)) \times D_p \quad (9)$$

4.3 IP address pricing (I)

For Networking, Microsoft Azure offers four types of deployments: Public IP Address (Static or Dynamic), Cloud Service VIP, Reserved IP Address and Instance-level Public IP Address (ILPIP). However this is a different model compared to AWS. Amazon Web Services do not charge any fees for IP addresses unless an IP address has been reserved but is not attached to an instance. This could be seen as a resource wastage fee. Much like Amazon Web Services, Google Cloud Platform do not charge the client any additional fees for IP addresses unless allocated to an unused server.

4.4 Support pricing (S)

AWS, Azure, and GCP offer a range of support plans, covering Basic, Developer, Business, and Enterprise with corresponding prices. Selecting an appropriate support package would be a business decision.

5 A worked example

In this section the cost model equations are applied to the frequency data of servers in use together with current (2018) pricing for both Azure, and AWS. For brevity this is limited to the subset of popular servers shown in Table 1. This covers 336 out of 506 systems in use in the data centre. Since these are all running Windows Server variants, the price includes any operating system licence fees. Both providers offer alternative one year, and three years payment plans for reserved instances. In this example the least advantageous one year payment discount has been chosen.

We obtain sample pricing vectors from [1], and [8]. Where no exact server match was present, the closest higher rated server (in terms of cores or ram) was substituted. This is an inevitable consequence of rapid technological progress, where the machine specification continually increases for the same price point.

The pricing vectors in (USD \$ per hour) for the server classes from Table 1 for AWS and Azure are then respectively:

$$\mathbf{p}_A = [0.0644, 0.0644, 0.1208, 0.2266, 0.354, 0.2266, 0.2266, 0.768, 0.768, 3.84]$$

$$\mathbf{p}_M = [0.065, 0.065, 0.096, 0.150, 0.345, 0.376, 0.376, 0.315, 0.495, 6.016]$$

5.1 AWS active servers

Assuming a mean discount rate of $D = 28\%$, then servers that are active for more than $730(1 - D) = 526$ hours per month (see 4.1) will be run as reserved instances. This includes 90 out of 336 servers. The frequency \mathbf{f}_R of reserved instances is then:

$$\mathbf{f}_R = 526 \times [9, 11, 16, 22, 5, 1, 8, 7, 11, 0]$$

$$C_R = \text{sum}(\mathbf{p}_A \cdot \mathbf{f}_R) = 13591.524$$

whilst standard instance hours are:

$$\mathbf{f}_S = [4983, 11033, 5947, 2009, 2208, 5873, 2032, 2003, 1895, 332]$$

$$C_S = \text{sum}(\mathbf{p}_A \cdot \mathbf{f}_S) = 9080.274$$

The cost of running these 336 servers with the known workload levels under AWS is then \$22671.80 per month.

5.2 Azure active servers

Assuming a mean discount rate of $D = 26\%$, then servers that are active for more than $730(1 - D) = 540$ hours per month are better reserved.

The frequency \mathbf{f} of reserved Instances is then:

$$\mathbf{f}_R = 540 \times [9, 11, 16, 22, 5, 1, 7, 7, 10, 0]$$

$$C_R = \text{sum}(\mathbf{p}_M \cdot \mathbf{f}_R) = 9750.984$$

whilst standard instance hours are:

$$\mathbf{f}_S = [4983, 11155, 5825, 2009, 2208, 5873, 2488, 2003, 2432, 332]$$

$$C_S = \text{sum}(\mathbf{p}_A \cdot \mathbf{f}_S) = 9721.944$$

The cost of running the sample server sets with their current workload levels under Microsoft Azure is \$19472.93 per month.

5.3 AWS IP address price (I)

As stated, active servers do not require an IP address. This is only needed on AWS when the server is not running and the address is so 246 IP addresses are required for AWS.

6 Conclusions and future work

This paper has described and motivated a cost model that allow a ready comparison of different cloud service providers when applied to an industrial scale computing workload. The equations have been evaluated against a dataset acquired from industrial activity. This evaluation shows that Cloud Computing costs may vary by up to 17%

when comparing compute service providers. It may also be observed that the $\approx 27\%$ of servers that would be better operated as reserved instances consume the majority of compute costs.

An observation from the case study is that workloads are cyclic, reflecting daily, weekly, and other periodic peaks in activity. Consequently there is scope for further mathematical sophistication. Finally, note that the cost model does not take into account charges for part hours or variable discount rates.

Another limitation of the study is that it only considers costs directly related to computing (i.e. CPU, Disk, Memory). Other cost savings that are factored into cloud prices include (but are not limited to), power costs, cooling, and premises. Additional expenses related to personnel and training are also excluded from this study, since these data are not readily available in a commercial environment.

In future work we are looking to address deficiencies in the current approach, that nevertheless has shown the financial advantage of cloud computing in an enterprise setting.

Acknowledgements Nanlin Jin is partly funded by Innovate UK—Knowledge Transfer Partnerships (Project ID: 10596).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

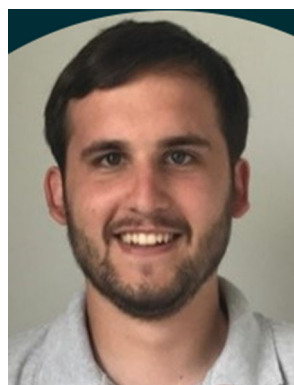
1. Amazon Web Services. (2018). AWS EC2 pricing. <http://aws.amazon.com/ec2/pricing>.
2. Bhardwaj, S., Jain, L., & Jain, S. (2010). Cloud computing: A study of infrastructure as a service (iaas). *International Journal of Engineering and Information Technology*, 2(1), 60–63.
3. Gallo, H., et al. (2016). AWS remains dominant despite microsoft and google growth surges. Synergy Research Group. <https://www.srgresearch.com/articles/aws-remains-dominant-despite-microsoft-and-google-growth-surges>.
4. Karidis, J., Moreira, J. E., & Moreno, J. (2009). True value: Assessing and optimizing the cost of computing at the data center level. In *Proceedings of the 6th ACM conference on computing frontiers* (pp. 185–192). ACM.
5. Khalid, A. (2010). Cloud computing: Applying issues in small business. In *International conference on signal acquisition and processing, 2010. ICSAP'10* (pp. 278–281). IEEE.
6. Li, A., Yang, X., Kandula, S., & Zhang, M. (2010). Cloudcmp: Comparing public cloud providers. In *Proceedings of the 10th ACM SIGCOMM conference on internet measurement* (pp. 1–14). ACM.
7. Mell, P., & Grance, T. (2010). The NIST definition of cloud computing. *Communications of the ACM*, 53(6), 50.
8. Microsoft. (2018). Windows virtual machines pricing. <https://azure.microsoft.com/en-gb/pricing/details/virtual-machines/windows/>.

9. Stamford, C. (2016). Gartner says by 2020 cloud shift will affect more than 1 trillion in it spending. <http://www.gartner.com/newsroom/id/3384720>.
10. Vella, E., Yang, L., Anwar, N., & Jin, N. (2018). Adoption of cloud computing in hotel industry as emerging services. In *Proceedings of the transforming digital worlds* (pp. 218–228). Springer International Publishing.
11. Vikas, S., Gurudatt, K., Vishnu, M., & Prashant, K. (2013). Private vs public cloud. *International Journal of Computer Science & Communication Networks*, 3(2), 79.
12. Weinman, J. (2014). The nuances of cloud economics. *IEEE Cloud Computing*, 1(4), 88–92.
13. Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud computing: State-of-the-art and research challenges. *Journal of Internet Services and Applications*, 1(1), 7–18.

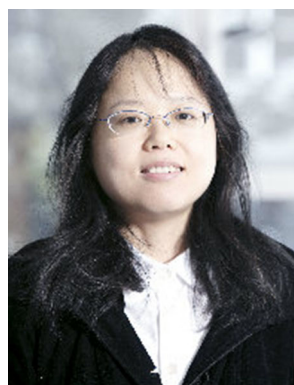


Dr Jeremy Ellman received the B.Sc. degree in experimental psychology from the University of Sussex, U.K., in 1978, the M.Sc. degree in computer studies from the University of Essex, U.K., in 1979, and the Ph.D. degree in computer science from the University of Sunderland, U.K., in 2000. He has been a Senior Lecturer in computer science with the University of Northumbria, U.K., since 2003. His research

works focus on the fields of computational linguistics, artificial intelligence, big data, and cloud computing.



Mr Nathan Lee was awarded a BSc degree in Information Technology Management for Business in 2016 from Northumbria University at Newcastle. Currently he is working for a large Engineering Manufacturer, in the North East of the UK.



Dr Nanlin Jin has worked for EPSRC and Technology Strategy Board's projects as a post-doc. She has published in the research areas of data analytics, IoT, and heuristic optimization. Since 2016, she has been the Lead Academic for an Innovate UK - Knowledge Transfer Partnership project to research data communication in Internet of Things (IoT).